Algorithmic and Combinatorial Methods to Discover Low Weight Pseudo-Codewords

Shashi Kiran Chilappagari[1] Bane Vasic[1] Mikhail Stepanov[2]
Michael Chertkov[3]

[1]Dept. of Elec. and Comp. Eng., University of Arizona, Tucson, AZ.

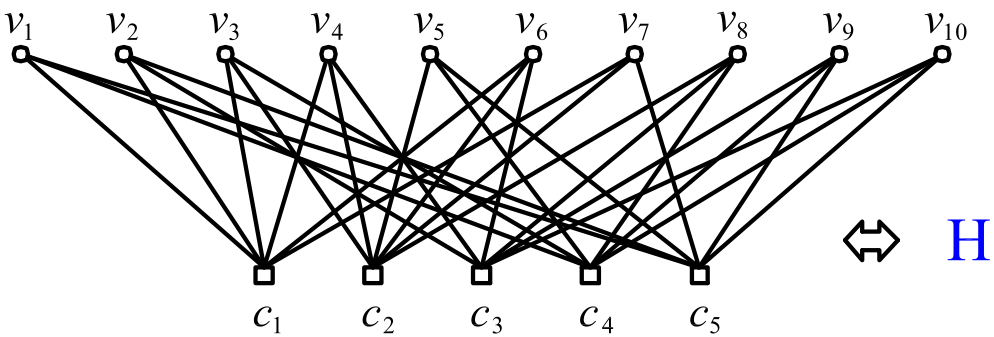
[2]Dept. of Mathematics, University of Arizona, Tucson, AZ.

[3]T-4, Theoretical Division, LANL, Los Alamos, NM.

Physics of Algorithms Workshop, Santa Fe, NM
September 2, 2009

Representation

- Codes based on sparse bipartite graphs
- Tanner graph with variable node set V and check node set C
- (d_v, d_c) regular bipartite graphs, irregular bipartite graphs



Tanner Graph

$$\Leftrightarrow \quad H = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

Decoding algorithms

Belief propagation (BP), Linear programming (LP) decoder, Gallager type decoders, Bit flipping decoders

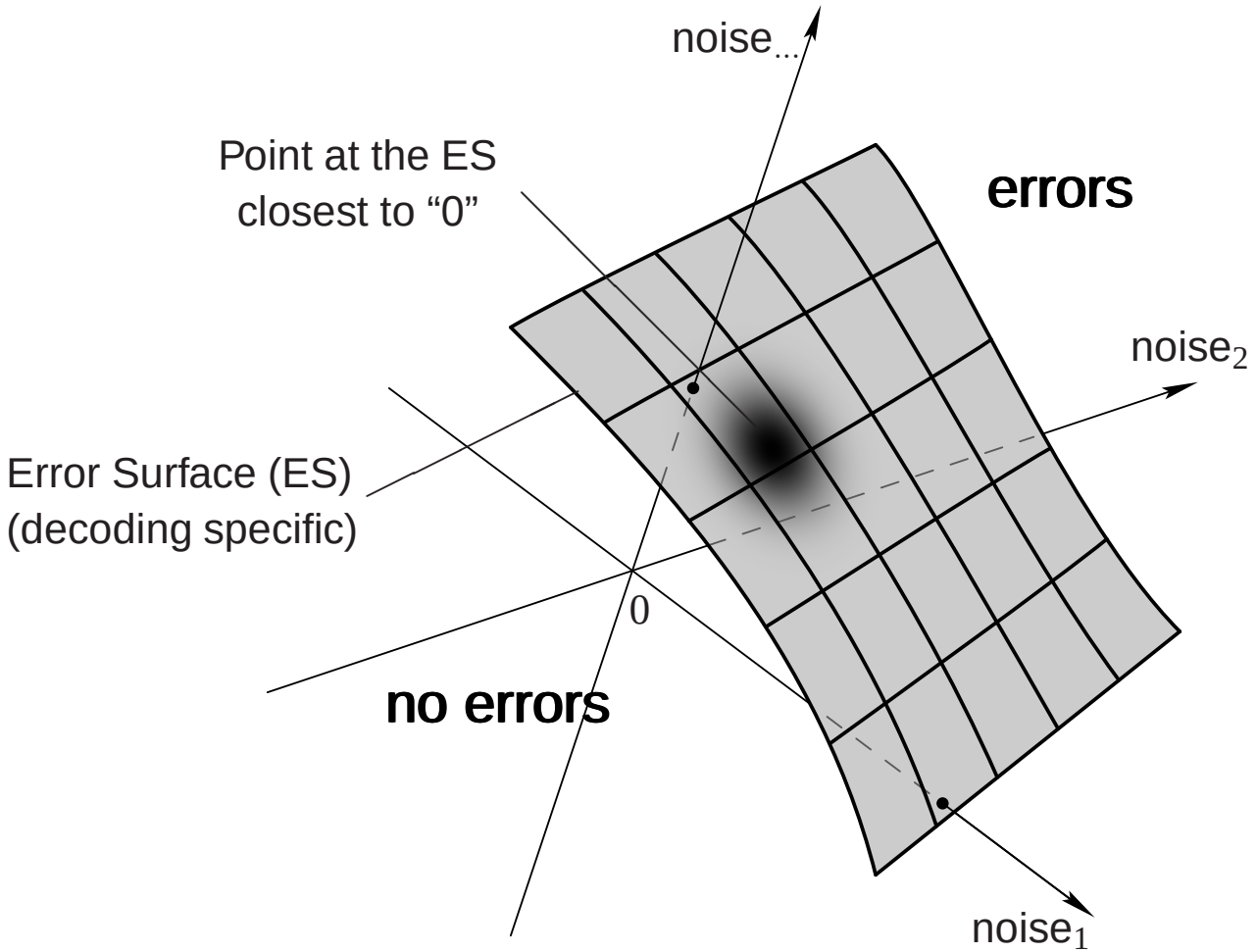
Asymptotic Analysis

Density evolution, EXIT charts, Expansion arguments (can be applied to finite length codes too)

Error Floor

- Abrupt degradation in performance in the high SNR region.
- Due to sub-optimality of the decoder.

Error Surface



Channel

- Codeword $\mathbf{y} = (y_1, \ldots, y_n)$ transmitted over a symmetric memoryless channel and received as $\hat{\mathbf{y}} = (\hat{y}_1, \ldots, \hat{y}_n)$.
- Log-likelihood-ratio (LLR) corresponding to variable node i

$$\gamma_i = \log\left(\frac{\Pr(\hat{y}_i|y_i = 0)}{\Pr(\hat{y}_i|y_i = 1)}\right).$$

ML-LP Decoder

- poly($\mathcal{C}$): Codeword polytope whose vertices correspond to codewords in $\mathcal{C}$
- Find $\mathbf{f} = (f_1, \ldots, f_n)$ minimizing the cost function $\sum_{i \in V} \gamma_i f_i$ subject to the constraint $\mathbf{f} \in \text{poly}(\mathcal{C})$

Realxed Polytope

Associating $(f_1, \ldots, f_n)$ with bits of the code we require

$$0 \leq f_i \leq 1, \quad \forall i \in V \tag{1}$$

For every check node j, let $N(j)$ denote the set of variable nodes which are neighbors of j. Let $E_j = \{T \subseteq N(j) : |T| \text{ is even}\}$. The polytope Q_j associated with the check node j is defined as the set of points $(\mathbf{f}, \mathbf{w})$ for which the following constraints hold

$$0 \leq w_{j,T} \leq 1, \qquad \forall T \in E_j \tag{2}$$

$$\sum_{T \in E_j} w_{j,T} = 1 \tag{3}$$

$$f_i = \sum_{T \in E_j, T \ni i} w_{j,T}, \quad \forall i \in N(j) \tag{4}$$

Now, let $Q = \cap_j Q_j$ be the set of points $(\mathbf{f}, \mathbf{w})$ such that (1)-(4) hold for all $j \in C$. The Linear Code Linear Program (LCLP) can be stated as

$$\min_{(\mathbf{f},\mathbf{w})} \sum_{i \in V} \gamma_i f_i, \text{ s.t. } (\mathbf{f}, \mathbf{w}) \in Q.$$

Features of LP Decoding

ML Certificate

- Integer solution is a codeword.
- If the LP decoder outputs a codeword, then the ML decoder would also output the same codeword.

Other Advantages of LP Decoding

- Discrete output: no numerical issues
- Attractive from analysis point of view
- Systematic sequential improvement

Assumptions

- Channel: binary symmetric channel (BSC)
- Transmission of the all-zero-codeword

Pseudo-codewords

A *pseudo-codeword*, $\mathbf{p} = (p_1, \ldots, p_n)$ where $0 \leq p_i \leq 1, \forall i \in V$, is simply equal to the output of the LP decoder.

Cost

- For the BSC, $\gamma_i = 1$ if $i = 0$ and $\gamma_i = -1$ if $i = 1$
- The output of the LP decoder on a received vector $\mathbf{r}$ is the pseudo-codeword $\mathbf{p}$ which has the minimum cost associated with it, where the $cost(\mathbf{r}, \mathbf{p})$ is given by

$$cost(\mathbf{r}, \mathbf{p}) = \sum_{i \notin \text{supp}(\mathbf{r})} p_i - \sum_{i \in \text{supp}(\mathbf{r})} p_i.$$

- The cost associated with decoding any vector $\mathbf{r}$ to the all-zero-codeword is zero

Example

- Consider a code of length 7
- Let $\mathbf{p} = (1, 0, \frac{1}{2}, \frac{2}{3}, \frac{1}{3}, 0, \frac{1}{2})$ be a pseudo-codeword
- Let the received vector be $\mathbf{r} = (1, 1, 0, 0, 0, 0, 0)$. Then, $\gamma = (-1, -1, 1, 1, 1, 1, 1)$
- $cost(\mathbf{r}, \mathbf{0}) = 0$ and $cost(\mathbf{r}, \mathbf{p}) = 1$
- If $\mathbf{r} = (1, 0, 0, 1, 0, 0, 0)$. Then, $\gamma = (-1, 1, 1, -1, 1, 1, 1)$
- $cost(\mathbf{r}, \mathbf{0}) = 0$ and $cost(\mathbf{r}, \mathbf{p}) = -\frac{1}{3}$

Weight and Median

Weight of a pseudo-codeword

Forney *et al.* '01

Let $\mathbf{p} = (p_1, \ldots, p_n)$ be a pseudo-codeword distinct from the all-zero-codeword. Let e be the smallest number such that the sum of the e largest p_is is at least $\left(\sum_{i \in V} p_i\right)/2$. Then, the BSC *pseudo-codeword weight* of $\mathbf{p}$ is

$$w_{BSC}(\mathbf{p}) = \begin{cases} 2e, & \text{if } \sum_e p_i = \left(\sum_{i \in V} p_i\right)/2; \\ 2e-1, & \text{if } \sum_e p_i > \left(\sum_{i \in V} p_i\right)/2. \end{cases}$$

Median

The median noise vector (or simply the median) $M(\mathbf{p})$ of a pseudo-codeword $\mathbf{p}$ distinct from the all-zero-codeword is a binary vector with support $S = \{i_1, i_2, \ldots, i_e\}$, such that $p_{i_1}, \ldots, p_{i_e}$ are the $e (= \lceil w_{BSC}(\mathbf{p})/2 \rceil)$ largest components of $\mathbf{p}$.

Instanton

The BSC *instanton* $\mathbf{i}$ is a binary vector with the following properties:

1. There exists a pseudo-codeword $\mathbf{p}$ such that $cost(\mathbf{i}, \mathbf{p}) \leq cost(\mathbf{i}, \mathbf{0}) = 0$
2. For any binary vector $\mathbf{r}$ such that $\text{supp}(\mathbf{r}) \subset \text{supp}(\mathbf{i})$, there exists no pseudo-codeword with $cost(\mathbf{r}, \mathbf{p}) \leq 0$. The size of an instanton is the cardinality of its support.

Theorem

Let $\mathbf{i}$ be an instanton. Then for any binary vector $\mathbf{r}$ such that $\text{supp}(\mathbf{i}) \subset \text{supp}(\mathbf{r})$, there exists a pseudo-codeword $\mathbf{p}$ satisfying $cost(\mathbf{r}, \mathbf{p}) \leq 0$.

Instanton Search Algorithm

Initialization ($l = 0$) step

Initialize to a binary input vector $\mathbf{r}$ containing sufficient number of flips so that the LP decoder decodes it into a pseudo-codeword different from the all-zero-codeword. Apply the LP decoder to $\mathbf{r}$ and denote the pseudo-codeword output of LP by $\mathbf{p}^1$.

$l \geq 1$ step

Take the pseudo-codeword $\mathbf{p}^l$ (output of the $(l-1)$ step) and calculate its median $M(\mathbf{p}^l)$. Apply the LP decoder to $M(\mathbf{p}^l)$ and denote the output by $\mathbf{p}_{M_l}$. Only two cases arise:

- $w_{BSC}(\mathbf{p}_{M_l}) < w_{BSC}(\mathbf{p}^l)$. Then $\mathbf{p}^{l+1} = \mathbf{p}_{M_l}$ becomes the l-th step output/$(l+1)$ step input.
- $w_{BSC}(\mathbf{p}_{M_l}) = w_{BSC}(\mathbf{p}^l)$. Let the support of $M(\mathbf{p}^l)$ be $S = \{i_1, \ldots, i_{k_l}\}$. Let $S_{i_t} = S \backslash \{i_t\}$ for some $i_t \in S$. Let $\mathbf{r}_{i_t}$ be a binary vector with support S_{i_t}. Apply the LP decoder to all $\mathbf{r}_{i_t}$ and denote the i_t-output by $\mathbf{p}_{i_t}$. If $\mathbf{p}_{i_t} = \mathbf{0}, \forall i_t$, then $M(\mathbf{p}^l)$ is the desired instanton and the algorithm halts. Else, $\mathbf{p}_{i_t} \neq \mathbf{0}$ becomes the l-th step output/$(l+1)$ step input.

Illustration of the ISA

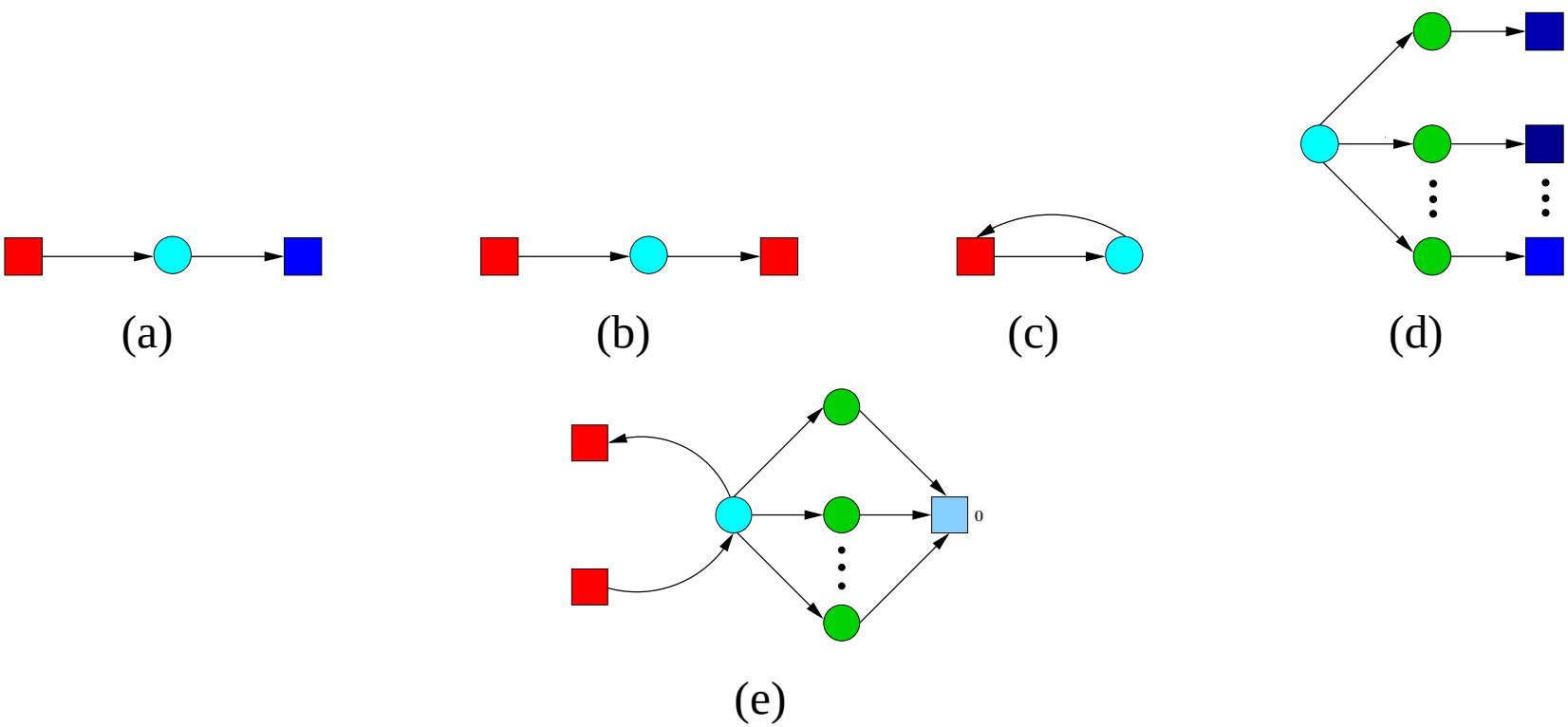


Figure: Squares represent pseudo-codewords and circles represent medians or related noise configurations (a) LP decodes median of a pseudo-codeword into another pseudo-codeword of smaller weight (b) LP decodes median of a pseudo-codeword into another pseudo-codeword of the same weight (c) LP decodes median of a pseudo-codeword into the same pseudo-codeword (d) Reduced subset (three different green circles) of a noise configuration (e.g. of a median from the previous step of the ISA) is decoded by the LP decoder into three different pseudo-codewords (e) LP decodes the median (blue circle) of a pseudo-codeword (low red square) into another pseudo-codeword of the same weigh (upper red square). Reduced subset of the median (three configurations depicted as green circles are all decoded by LP into all-zero-codeword. Thus, the median is an instanton.

The ISA Converges

Lemma 1

Let $M(\mathbf{p})$ be a median of $\mathbf{p}$ with support S. Then the result of LP decoding of any binary vector with support $S' \subset S$ and $|S'| < |S|$ is distinct from $\mathbf{p}$.

Lemma 2

If the output of the LP decoder on $M(\mathbf{p})$ is a pseudo-codeword $\mathbf{p}_M \neq \mathbf{p}$, then $w_{BSC}(\mathbf{p}_M) \leq w_{BSC}(\mathbf{p})$. Also, $cost(M(\mathbf{p}), \mathbf{p}_M) \leq cost(M(\mathbf{p}), \mathbf{p})$.

Theorem

$w_{BSC}(\mathbf{p}^l)$ and $|supp(M(\mathbf{p}^l))|$ are monotonically decreasing. Also, the ISA terminates in at most $2k_0$ steps, where k_0 is the number of flips in the input.

Instanton Statistics

Table: Instanton statistics obtained by running the ISA with 20 random flips for 10000 initiations for the Tanner code and the MacKay code

Code		Number of instantons of weight					
		4	5	6	7	8	9
Tanner code	Total		3506	1049	1235	1145	1457
	Unique		155	675	1028	1129	1453
MacKay code	Total	213	749	2054	2906	2418	1168
	Unique	26	239	1695	2864	2417	1168

Comments

Why efficient?

- Extremely efficient compared to naive search methods
- Brute force search, searching for instantons in bigger error patterns

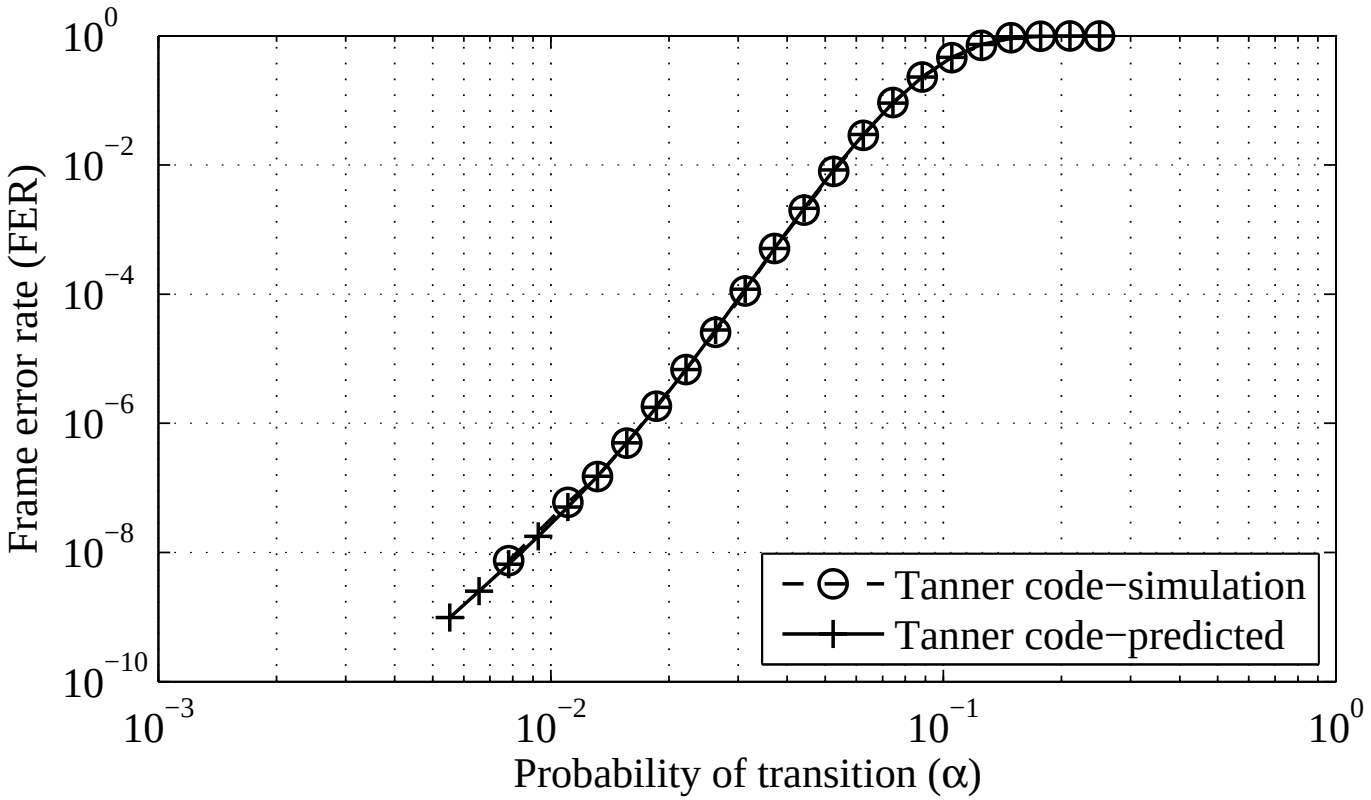
Why LP?

- Median difficult to define for other decoders
- No monotonic property for iterative decoders

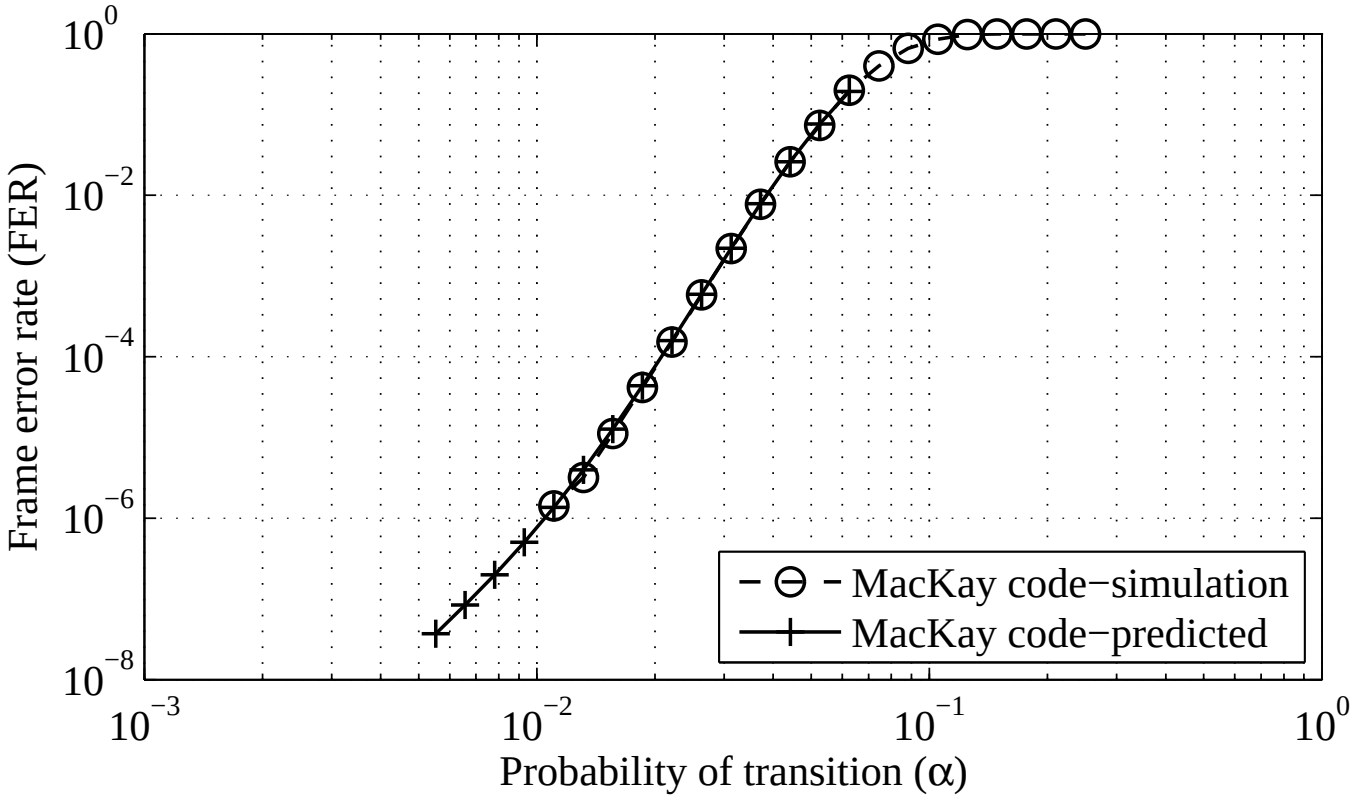
When to stop?

- Number of instantons grows with size
- To count N instantons, we need approximately $N \log(N)$ trials of ISA

The Tanner Code



MacKay Random Code



Trapping Sets for Iterative Decoders

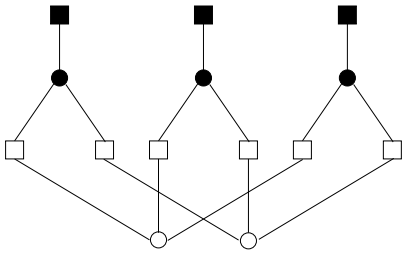
Theorem

Let $\mathcal{C}$ be an LDPC code with $d_v = 3$ left-regular Tanner graph G with girth $g \geq 10$. Then G always has a trapping set of size $g/2$ and no trapping sets of size less than $g/2$. Moreover, every shortest cycle is a trapping set.

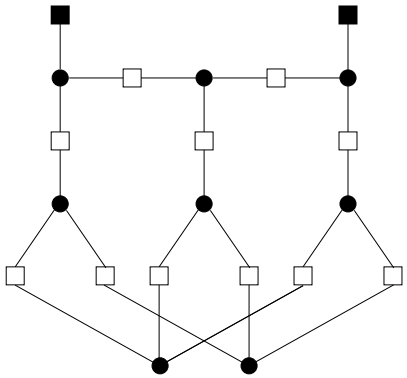
Theorem

Let $\mathcal{C}$ be an LDPC code with d_v-left-regular Tanner graph G. Let $\mathcal{T}$ be a set consisting of V variable nodes with induced subgraph $\mathcal{I}$. Let the checks in $\mathcal{I}$ be partitioned into two disjoint subsets; $\mathcal{O}$ consisting of checks with odd degree and $\mathcal{E}$ consisting of checks with even degree. Then $\mathcal{T}$ is a fixed set for the bit flipping algorithm (serial or parallel) iff : (a) Every variable node in $\mathcal{I}$ has at least $\lceil d_v/2 \rceil$ neighbors in $\mathcal{E}$, and (b) No $\lfloor d_v/2 \rfloor + 1$ checks of $\mathcal{O}$ share a neighbor outside $\mathcal{I}$.

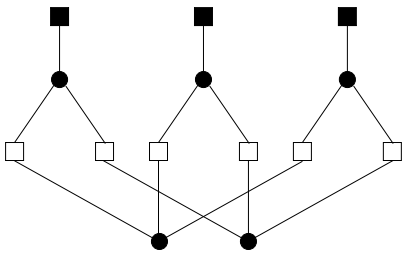
Relation Between Failures of Different Decoders



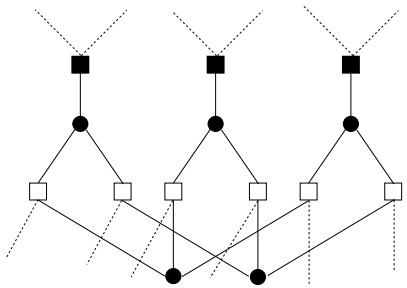
(a) A (5,3) trapping set for Gallager A algorithm.



(b) An (8,2) trapping set for iterative decoding over the AWGNC.



(c) The support of an instanton of size 5 for LP decoding over the BSC.



(d) The support of the lowest weight pseudo-codeword for LP decoding over the AWGN.

Concluding Remarks

- Use knowledge of failures of one decoder to find failures of others
  - Relation to trapping sets
  - Valid hyperflows
- Improving the code for one decoder potentially improves it for other decoders
- Construct codes avoiding known harmful configurations
- Tighten the LP decoder using the knowledge of instantons

References

- C. Di, D. Proietti, T. Richardson, E. Telatar, and R. Urbanke, "Finite length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1570–1579, June 2002.
- T. J. Richardson, "Error floors of LDPC codes," in *41st Annual Allerton Conf. on Communications, Control and Computing*, pp. 1426–1435, 2003.
- M. G. Stepanov, V. Chernyak, M. Chertkov, and B. Vasic, "Diagnosis of weaknesses in modern error correction codes: A physics approach," *Phys. Rev. Lett.*, vol. 95, p. 228701, Nov. 2005.
- N. Wiberg, *Codes and decoding on general graphs*. Ph.d., Univ. Linköping, Sweden, Dept. Elec. Eng., 1996.
- P. O. Vontobel and R. Koetter, "Graph-cover decoding and finite length analysis of message-passing iterative decoding of LDPC codes," Dec. 2005.
- J. Feldman, M. Wainwright, and D. Karger, "Using linear programming to decode binary linear codes," *Information Theory, IEEE Transactions on*, vol. 51, pp. 954–972, March 2005.
- M. Chertkov and M. Stepanov, "An efficient pseudocodeword search algorithm for linear programming decoding of LDPC codes," *Information Theory, IEEE Transactions on*, vol. 54, pp. 1514–1520, April 2008.
- S. K. Chilappagari, M. Chertkov, and B. Vasic, "Provably efficient instanton search algorithm for LP decoding of LDPC codes over the BSC," 2008. Submitted to *IEEE Trans. Inform. Theory*.